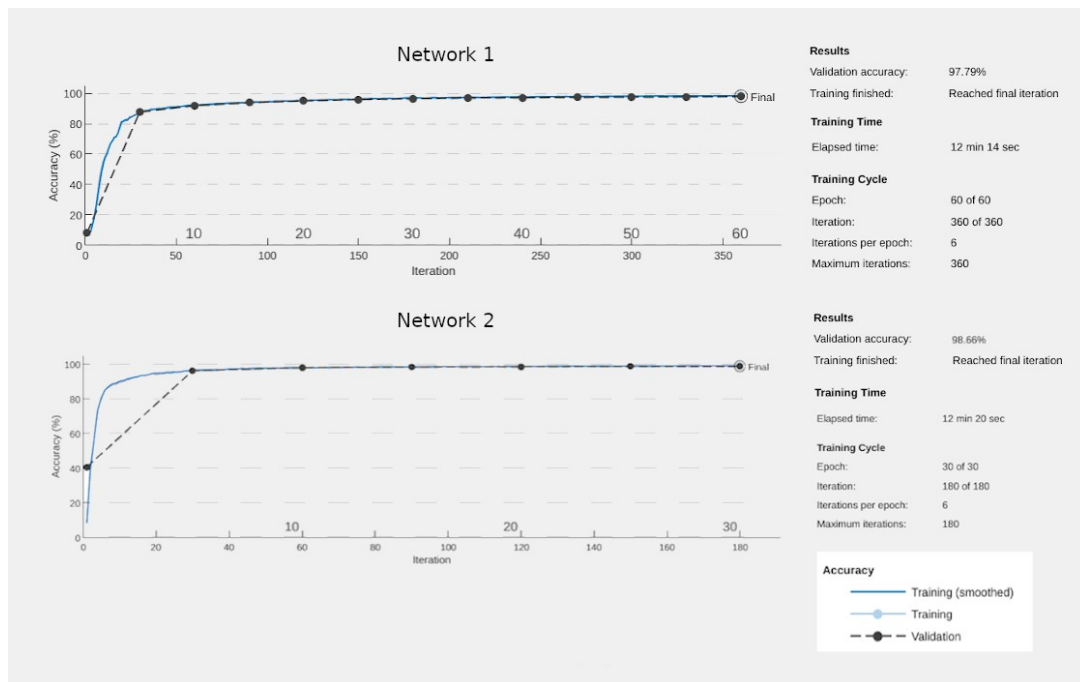


Homework 3

Problem 1: Convolutional networks

Following the instructions given we construct two networks and use data from the MNIST database to train them. The training progress for both networks is show in the following figure:



We can see that the second network reaches higher validation accuracy faster than the first network and the overall validation accuracy of the second network is better. We notice the same when we calculate the errors obtained on the training, validation and test sets:

	Network 1	Network 2
Training Set Error	0.0162	0.0094
Validation Set Error	0.0221	0.0134
Test Set Error	0.0192	0.0120

The second network performs better than the first because it is deeper and this way the multiple layers can learn more intermediate features between the raw data and the high-level classification. For example the first layer will train itself to recognize very basic things like edges while the next layer will recognise collection of edges such as shapes. Thus deeper layers will learn more complicated and higher-order features.